

AIX Sıkıştırma Yöntemleri

Yapılan testleri özetlersem.

Öncelikle AIX'e özel sıkıştırma programı bulmak güç. Lzop'u derleyerek çalıştırılabilir programı oluşturabildim.

Yada nadir olsa derlenmişide bulunabiliyor.

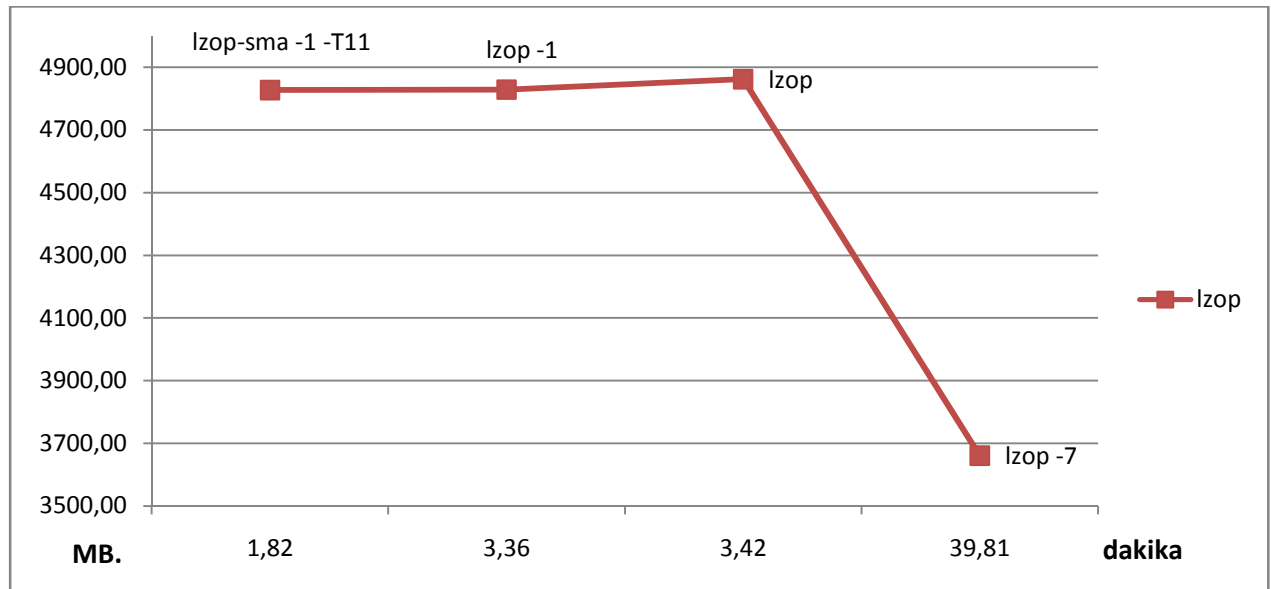
Lzop'un smp versiyonu var. Paralel şekilde sıkıştırma prosesi yaparak cpu'yu daha verimli kullanmayı hedefliyor.

Oluşturduğu dosya boyutunda fark yok ancak sıkıştırma süresi gerçekten çok hızlı.

Sıkıştırdığı boyut ise genel olarak kabul edilebilir ölçüde.

Çalıştırılan - Sıkıştırma İşlemi	Boyut (MB)	Süre (s)	Sıkış. oranı	Sıkıştırılan veri (MB/s)	Performans (Sıkış. oranı X işlenen veri)
Düz kopyalama	16029,9	175,70	1,00	91,23	91,23
lzop -1 test.dmp	4829,11	198,53	3,32	80,74	268,02
lzop -2 test.dmp (-2,3,4,5,6 aynı sonucu verir)	4862,21	205,08	3,30	78,17	257,70
lzop -7 test.dmp	3660,73	2388,60	4,38	6,71	29,39
gzip -1 test.dmp	3824,8	631,66	4,19	25,38	106,36
gzip -2 test.dmp	3719,86	627,46	4,31	25,55	110,09
gzip -3 test.dmp	3606,43	711,46	4,44	22,53	100,15
gzip -4 test.dmp	3381,14	830,86	4,74	19,29	91,47
gzip -5 test.dmp	3206,45	977,80	5,00	16,39	81,96
gzip test.dmp (-6 aynı sonucu verir)	3137,06	1308,53	5,11	12,25	62,60
lzop-sma -1 -T11 test.dmp	4827,65	109	3,32	147,06	488,31

LZOP'a genel bakış



Lzop'un en efektif kullanımı -1 yani fast parametresiyle gerçekleşiyor. Hem boyut hem de sıkıştırılmış veri oluşturma süresi az çıkıyor.

Lzop'u paralel kullnarak 16GB'lık bir dosyayı 1.82 dakikada (109 saniye) de oluşturabiliriz.

Bu şu anlama geliyor prod2 makinesinde 30 dakikada süren gzip -2 işlemi lzop -1 kullanarak 9.5 dakikada

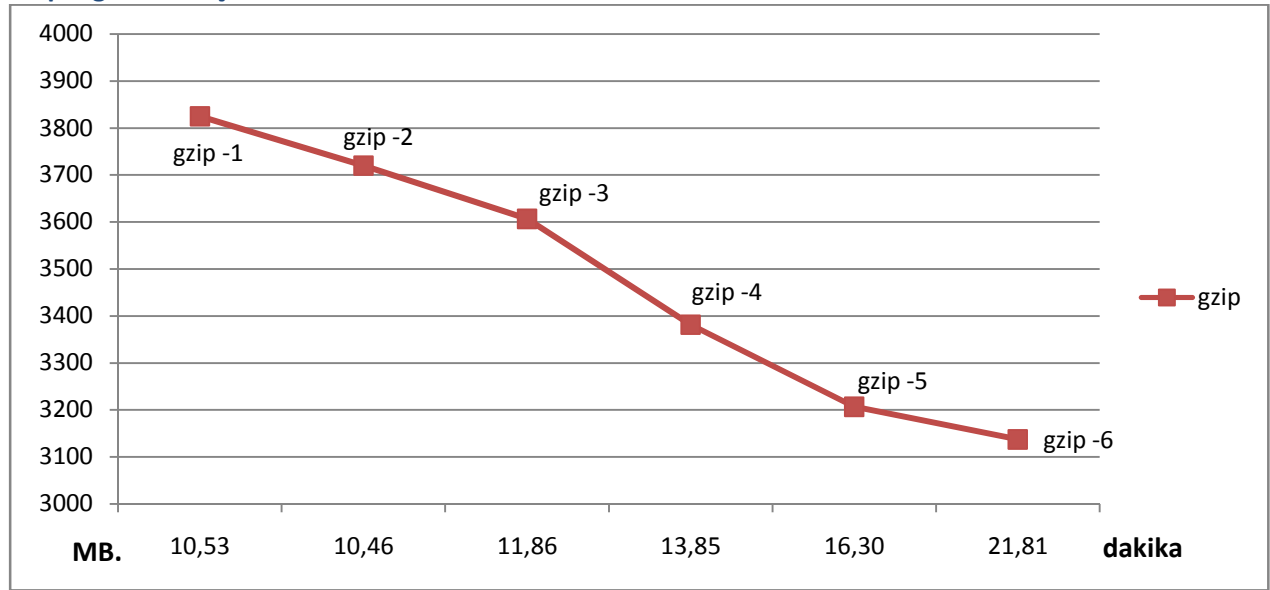
yine benzer şekilde lzop-sma -1 kullanarak sıkıştırma işlemini 5.2 dakikada tamamlayabileceğimiz anlamına geliyor.

Lzop sıkıştırma zamanında %25-%50 arasında kullanırken işlemciyi, lzop-sma %50-%75 arasında kullanmaktaydı.

Lzop-sma da bulunan T parametresiyle thread sayısı artırılabilir (8 ile 128 arasında)

Özellikle lzop-sma'nın dosyayı farklı bir isimle kopyalama hızından daha hızlı şekilde sıkıştırılmış dosya oluşturması beni şaşırttı.

Gzip'e genel bakış



Çeşitli dosyaların tahmini boyutunu sıkıştırma oranı vasıtasıyla hesap edebilirsiniz.

Örneğin gzip -5 ile 100GB lık bir data $100/5=20$ GB boyutuna inecektir.

Standart şekilde gzip işlemi yaklaşık $100/5.11=19.5$ GB civarında yer kaplayacaktır.

Lzop -1 de ise yine 100GB lık bir data diskte $100/3.32=30.1$ GB boyutunda yer kaplayacaktır.

Sıkıştırılmış dosyaları açmaya gelirse dikkat etmemiz alan performans alanından ziyade süre olmalıdır.

En başarılı sonucu en çabuk sürede sıkıştırılmış dosyayı açan lzop öne çıkıyor yine.

Çalıştırılan - Açma işlemi	Boyut (MB)	Süre (s)	Açma oranı	Açılan veri (MB/s)	Performans (Açma oranı X işlenen veri)
Düz Kopyalama	16029,9	175,70	1,00	91,23	91,23
lzop test.dmp (varsayılan 3)	4862,21	151,50	3,30	105,81	348,83
lzop -d test.dmp.lzo (lzop -1 den dönüş)	4829,11	149,81	3,32	107,00	355,18
lzop -d test.dmp.lzo (lzop -2 den dönüş)	4862,21	150,59	3,30	106,45	350,94

lzop -d test.dmp.lzo (lzop -7 den dönüş)	3824,8	245,95	4,19	65,18	273,15
gunzip test.dmp.gz (gzip -1 den dönüş)	3824,8	245,95	4,19	65,18	273,15
gunzip test.dmp.gz (gzip -2 den dönüş)	3719,86	236,32	4,31	67,83	292,30
gunzip test.dmp.gz (gzip -3 den dönüş)	3606,43	232,79	4,44	68,86	306,07
gunzip test.dmp.gz (gzip -4 ten dönüş)	3381,14	238,15	4,74	67,31	319,12
gunzip test.dmp.gz (gzip -5 ten dönüş)	3206,45	240,52	5,00	66,65	333,19
gunzip test.dmp (gzip den dönüş)	3137,06	224,52	5,11	71,40	364,83

Aşağıdakilerde kendi bilgisayarımda yaptığım testleri içeriyor.

Çalıştırılan - Sıkıştırma İşlemi	Boyut (MB)	Süre (s)	Sıkıştırma oranı	Sıkıştırılan veri (MB/s)	Performans (Sıkıştırma oranı x İşlenen veri)
Düz kopyalama	1760,68	165,00	1,00	10,67	10,67
arc -1	393,85	112,71	4,47	15,62	69,84
arc -2	269,28	117,59	6,54	14,97	97,90
arc -3	223,52	372,88	7,88	4,72	37,19
gzip -2	425,17	120	4,14	14,67	60,76
gzip -6	360,07	229	4,89	7,69	37,60
lzop -1	547,60	101,86	3,22	17,29	55,58
bzip2 -1	284,16	421,26	6,20	4,18	25,90

Kendi bilgisayarımda en iyi sonucu "freearc" ile elde ettim.

16GB lık datayı 20 dakikada (1203.91 saniye) de 2GB'a indirecek şekilde sıkıştırdı.

C:\e>arc a test.arc2 -m2 test.dmp

ARC 0.40 creating archive: test.arc2

Compressed 1 file, 16.808.566.784 => 2.147.725.888 bytes. Ratio 12.7%

Compression time 646.69 secs, speed 25.992 kb/s. Total 1203.91 secs

All OK

C:\e>arc a ufak.arc2 -m2 ufak.dmp

ARC 0.40 updating archive: ufak.arc2

Compressed 1 file, 1.846.210.560 => 282.365.187 bytes. Ratio 15.2%

Compression time 85.02 secs, speed 21.716 kb/s. Total 118.28 secs

All OK

Ayrıca 1760MB lık dosyayı paq9a'nın -1 parametresiyle 151MB -9 parametresiyle 72MB'e indirmek mümkün.

Mükemmel bir sıkıştırma sunuyor ancak çok yavaş. İşlem saatler sürdü.

Hem hızlı hemde gzip'ten daha iyi bir sıkıştırma sunuyor. Maalesef aix için uygun versiyonunu bulamadığım için aix'te çalıştırma olanağım olmadı.

Test değerlerine gelirse

Yine de yaklaşık %2 lik bir hata payı ile gerçek sonuca yakın bir değer çıkaracağını düşünüyorum.

Tüm deęerler tahmini olduęunu unutmamak lazım, işlemci + disk kullanımının kullanımda olup olmadığına göre ve sıkıştırılacak veriye göre deęişecektir.

Hata olmasını engellemek amaçlı aynı sonuçları 1 den fazla tekrarlayıp ortalamasını aldım.

Sonuç olarak eęer amacımız datayı "kabul" edilebilir şekilde ve hızlı şekilde sıkıştırmaksa LZOP (<http://www.oberhumer.com>, <http://www.lzop.de/>) en hızlısı.

Daha iyi şekilde ve yine hızlı şekilde sıkıştırmak istersek FreeARC (<http://freearc.org/>). Bence kullanılabilir en iyi sıkıştırma programı.

Zaman sıkıntımız yoksa PAQ sıkıştırma konusunda harikalar yaratıyor (<http://cs.fit.edu/~mmahoney/compression/paq.html>)

Kaynaklar:

http://www.maximumcompression.com/data/summary_mf2.php#data

<http://lemley.net/smp.html>